

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2001, 16:12:19 ; Search time 31.1 Seconds
(without alignments)
1711.009 Million cell updates/sec

Title: US-09-455-486-6
Perfect score: 2351
Sequence: 1 MESISMWGPXSLSEITCLPN.....ALVPSIVILDLQLCRYPD 454

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL.15:**

- 1: sp_archaea:**
- 2: sp_bacteria:**
- 3: sp_fungi:**
- 4: sp_human:**
- 5: sp_invertebrate:**
- 6: sp_mammal:**
- 7: sp_mhc:**
- 8: sp_organelle:**
- 9: sp_phage:**
- 10: sp_plant:**
- 11: sp_rodent:**
- 12: sp_virus:**
- 13: sp_vertebrate:**
- 14: sp_unclassified:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	ID	Description
1	1269	54.0	4 Q9NVB5	Q9nvb5 homo sapien
2	717	30.5	4 Q9UHE8	Q9uhe8 homo sapien
3	717	30.5	4 Q95034	Q95034 homo sapien
4	577	24.5	4 Q9Y6U5	Q9y6u5 homo sapien
5	156.5	6.7	239 Q9KKR6	Q9kr6 streptomyc
6	147	6.3	213 Q29059	Q29059 archaeoglob
7	142	6.0	212 Q29370	Q29370 archaeoglob
8	142	6.0	224 Q59661	Q59661 methanobact
9	136.5	5.8	217 Q3JN78	Q3jn78 streptomyc
10	119.5	5.1	216 Q50637	Q50637 leptospira
11	116.5	5.0	242 Q3PCP1	Q3pcp1 xylella fas
12	114	4.8	207 Q3RIU0	Q3riu0 streptomyc
13	112.5	4.8	207 Q3RIT8	Q3rit8 streptomyc
14	112	4.8	408 Q9V8W8	Q9v8w8 drosophila
15	111.5	4.7	1228 Q12754	Q12754 saccharomyc
16	109	4.6	219 Q3RIU2	Q3riu2 streptomyc
17	109	4.6	320 Q9YVM4	Q9yvm4 melanopius
18	109	4.6	1301 Q23530	Q23530 arabidopsis
19	105.5	4.5	348 Q79242	Q79242 poecilia ca

20	105	4.5	442	2	Q25240	Q25240 helicobacte
21	105	4.5	604	1	Q93700	Q93700 sulfolobus
22	102.5	4.4	391	8	Q9TKV6	Q9tkv6 nephroselmi
23	102.5	4.4	1242	3	Q60103	Q60103 schizosacch
24	102	4.3	574	3	Q74441	Q74441 schizosacch
25	101	4.3	320	1	O58066	O58066 pyrococcus
26	101	4.3	348	8	O47399	O47399 poecilia pe
27	101	4.3	420	1	O26447	O26447 methanobact
28	101	4.3	476	8	Q35809	Q35809 saccharomyc
29	101	4.3	501	10	O80695	O80695 arabidopsis
30	101	4.3	503	10	Q9SAB0	Q9sab0 arabidopsis
31	100.5	4.3	346	8	Q9XLI5	Q9xli5 smithornis
32	100.5	4.3	1780	10	Q9ZT82	Q9zt82 arabidopsis
33	100	4.3	348	8	Q9ZT30	Q9zt30 rivulus obs
34	100	4.3	395	1	O37766	O37766 pyrococcus
35	100	4.3	3411	12	Q9YRV3	Q9yrv3 yellow feve
36	99.5	4.2	348	8	O47406	O47406 poecilia he
37	99.5	4.2	348	8	O79244	O79244 poecilia la
38	99.5	4.2	348	8	O79246	O79246 poecilia bu
39	99.5	4.2	919	5	Q9VCV7	Q9vcv7 drosophila
40	99.5	4.2	1518	5	Q9NGP5	Q9ngp5 dictyosteli
41	99	4.2	389	10	Q9SX90	Q9sxx90 arabidopsis
42	99	4.2	444	2	Q9ZFT3	Q9zft3 streptococc
43	98.5	4.2	322	2	Q9XCR5	Q9xcr5 enterococcu
44	98.5	4.2	323	2	Q9XBX3	Q9xbx3 enterococcu
45	98.5	4.2	323	2	Q9S3S8	Q9s3s8 enterococcu

ALIGNMENTS

RESULT 1

Q9NVB5 PRELIMINARY: PRT: 488 AA.
AC Q9NVB5;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE CDNA FLJ10829 FIS, CLONE NT2RP4001138.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isoqai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA Nakamura Y., Nagahari K., Masuho Y., Sasaki N.;
RT "NEDO human cDNA sequencing project."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK001691; BAA91839.1; -.
SQ SEQUENCE 488 AA; 54616 MW; BC0BCA483335AAD6 CRC64;

! Query Match 54.0%; Score 1269; DB 4; Length 488;
Best Local Similarity 54.3%; Pred. No. 3.4e-89;
Matches 241; Conservative 83; Mismatches 108; Indels 12; Gaps 4;

QY	3	SISMWGPXSLSEITCLPN	INGIKDARKVTGVVIGSGDPAKSLTIRLCRGTHVVGSRN	62
Db	11	SHLVDSDSLAK--VPD-----EAPK--VGILGSGDFARSLATRLVGSGFKVVGSRN	60	
QY	63	PKFASEFFPHVVDVTHEDALTKNIFVAIHREHYVTSIWLDRHLVLVGLKILLDVSN--NM	120	
Db	61	PRTRALYPSAAQVTFQEEAVSSPEVIFVAFREHYSSLCISLSDQLAGKILLDVSNPTQ	120	
QY	121	RINQYPSNAEYLAFLPDSLIKGFNVVSAVALQPKDASQVVCISNNTOARQVTE	180	
Db	121	EHLQHRSENAEYLAFLPCTVVKAFNVISANTLQNGPRDGRNQVPCIDQGEAKRATSE	180	
QY	181	LARQLNFIPIDLGSLSSAREIENLRLRLFTLWRGPPVVVAISLATFFFLYSFVRDVIHYA	240	
Db	181	LAQLNFIPIDLGSLSSAREIENLRLRLFTLWRGPPVVVAISLATFFFLYSFVRDVIHYA	240	

Db 181 MALAWGFMPVDMGSLASAWVEAMPLRLPAWKVPTLLALGLFVCFYAYNFVRDLQPVV 240
 Qy 241 RNOQSDFYKIEIVNKTLPVAVITLLSLVYLAGLLAAAYOLYGTGYRRPFPWLEWLQ 300
 Db 241 QESQNKFKLPVSVNTTLPVAVVLSLVLPGLAALQLRGTYKQRFDPDLHWLQ 300
 Qy 301 CRKQGLLSFFFAVHVAVYSLCLPMRSEYRFLNMAVQQVQHANIENSWNEEYVRIEM 360
 Db 301 HRKQIGLLSFFCAALHALYSCLPLRRHRYDLVNLAVKQVLANKSLHWEEVWRMEIY 360
 Qy 361 ISFGIMSLGILLSLAVTSIPSVNALNWRSEFSTQSTLGYVALLISTFHVLIYWKRAFE 420
 Db 361 LSLGLVALGTLUSLAVTSIPSVNALNWRSEFVQSSLGFLVALVLTHTLTLYGWTRAFE 420
 Qy 421 EBYRFRYTPPNFVALVPSVIL 444
 Db 421 ESRKFKYLPPTFTLLVPCVIL 444

RESULT 2

ID Q9UHE8 PRELIMINARY; PRT; 339 AA.
 AC Q9UHE8
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE SIX TRANSMEMBRANE EPITHELIAL ANTIGEN OF PROSTATE.
 GN STEAP1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hubert R.S., Vivanco I., Chen E., Rastegar S., Leong K.,
 RA Mitchell S.C., Madraswala R., Zhou Y., Kuo J., Raitano A.B.,
 RA Jakobovits A., Safiran D.C., Afar D.E.H.;
 RT "STEAP: A prostate-specific cell-surface antigen highly expressed in
 RT human prostate tumors."
 RL Proc. Natl. Acad. Sci. U.S.A. 96:14523-14528(1999).
 DR EMBL: AF186249; AAF17479.1;
 KW Transmembrane.
 SQ SEQUENCE 339 AA; 39851 MW; 55443AI70C870387 CRC64;

Query Match 30.5%; Score 717; DB 4; Length 339;
 Best Local Similarity 54.9%; Pred. No. 3.7e-47;
 Matches 130; Conservative 48; Mismatches 59; Indels 0; Gaps 0;

Qy 208 LFTLWRGPVVVAISLATFFFLYSFVRDVIHPYARNQSDFYKIEIVNKTLPVAVITLL 267
 Db 67 LFPQWHLPIKIAAIIASLTFLYTLREVHPLATSHQOYFYKIPILVINKVLPWVSITLL 126
 Qy 268 SLVYLAGLLAAAYOLYGTGYRRPFPWLEWLQCRKQLGSLFFFAVHVAVYSLCLPMR 327
 Db 127 ALVLPGVIAAIVQLHNGTKYKKEPPLDKWMLTRKQGLSFFFAVHAIYSYSPMR 186
 Qy 328 SERYFLNMAVQQVQHANIENSWNEEYVRIEMYSFGIMSLGILLSLAVTSIPSVNALN 387
 Db 187 SYRKLLNWAYQQVQNKEDAWIEHDVWRMEIYVSLGIVGLAILALLAVTSIPSVDSLT 246
 Qy 388 WREFSFTOSTLGYVALLISTFHVLIYWKRAFEYRFTYPPNFVALVLPISVIL 444
 Db 247 WREFHYIQSKLGIVSLGLTTHALIFAWNKWIDIKQFWWYTPPTFMIAVFLPIVILI 303

RESULT 3

ID O95034 PRELIMINARY; PRT; 373 AA.
 AC O95034
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-AUG-1999 (TrEMBLrel. 11, Last annotation update)

DE WUGSC:H_RG041D11.1 PROTEIN (WUGSC:H_DJ1121E10.1 PROTEIN) (FRAGMENT)
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Abu-Threideh J., Stoneking T., Langston Y., Maupin R.;
 RT "The sequence of Homo sapiens BAC clone RG041D11."
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Waterston R.H.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Waterston R.;
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 25-373 FROM N.A.
 RX MEDLINE=99063792; PubMed=9847074;
 RA Sulston J.E., Waterston R.;
 RT "Toward a complete human genome sequence."
 RL Genome Res. 8:1097-1108(1998).
 RN [5]
 RP SEQUENCE OF 25-373 FROM N.A.
 RA Kraemer J., Miller N., Ozersky P., Walker C.;
 RT "The sequence of Homo sapiens PAC clone DJ1121E10."
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE OF 25-373 FROM N.A.
 RA Waterston R.H.;
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE OF 25-373 FROM N.A.
 RA Waterston R.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC005053; AAC79150.1;
 DR EMBL: AC004969; AAD15620.1;
 FT NON_TER 1
 SQ SEQUENCE 373 AA; 43574 MW; B8BE3C86F903F775 CRC64;

Query Match 30.5%; Score 717; DB 4; Length 373;
 Best Local Similarity 54.9%; Pred. No. 4.2e-47;
 Matches 130; Conservative 48; Mismatches 59; Indels 0; Gaps 0;

Qy 208 LFTLWRGPVVVAISLATFFFLYSFVRDVIHPYARNQSDFYKIEIVNKTLPVAVITLL 267
 Db 101 LFPQWHLPIKIAAIIASLTFLYTLREVHPLATSHQOYFYKIPILVINKVLPWVSITLL 160
 Qy 268 SLVYLAGLLAAAYOLYGTGYRRPFPWLEWLQCRKQLGSLFFFAVHVAVYSLCLPMR 327
 Db 161 ALVLPGVIAAIVQLHNGTKYKKEPPLDKWMLTRKQGLSFFFAVHAIYSYSPMR 220
 Qy 328 SERYFLNMAVQQVQHANIENSWNEEYVRIEMYSFGIMSLGILLSLAVTSIPSVNALN 387
 Db 221 SYRKLLNWAYQQVQNKEDAWIEHDVWRMEIYVSLGIVGLAILALLAVTSIPSVDSLT 280
 Qy 388 WREFSFTOSTLGYVALLISTFHVLIYWKRAFEYRFTYPPNFVALVLPISVIL 444
 Db 281 WREFHYIQSKLGIVSLGLTTHALIFAWNKWIDIKQFWWYTPPTFMIAVFLPIVILI 337

RESULT 4

Q9Y6U5 PRELIMINARY; PRT; 264 AA.
 ID Q9Y6U5
 AC Q9Y6U5
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
 DE WUGSC:H_RG087E15.1 PROTEIN (FRAGMENT).
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99063792; PubMed=9847074;
 RA Sulston J.E., Waterston R.;
 RT "Toward a complete human genome sequence.";
 RL Genome Res. 8:1097-1108(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Strong C., Layman D., Graves T., Strowmatt C.;
 RT "The sequence of Homo sapiens BAC clone RG087E15.";
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Waterston R.H.;
 RT Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Waterston R.H.;
 RT Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Waterston R.;
 RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC005061; AAD43182.1;
 FT NON_TER 1 264
 FT NON_TER 264 264
 SQ SEQUENCE 264 AA; 31203 MW; 35C9483003557E72 CRC64;

Query Match 24.5%; Score 577; DB 4; Length 264;
 Best Local Similarity 56.4%; Pred. No. 1.4e-36;
 Matches 106; Conservative 36; Mismatches 46; Indels 0; Gaps 0;

QY 208 LFTLRGPPVVAISLATEFFLYSFYRDVIHPYARNOQSDFYKIPIETIVNKTLPVIAITLL 267
 DB 77 LPQHLPIKIAAVMSLTFLLREVIHPILATSHQYFYKIPILVINKVLPVMSITLL 136
 QY 268 SLVYLGLAAAYQLYGKYPFRPPWLTWLCQKRLGLSFFFAVMHVAYSCLPMMR 327
 DB 137 ALVYLPVIAAVQVNGTKYKFFHDLKMLTKQGLSLSLFFAVLHAIYTLSYAMR 196
 QY 328 SERYLFLNMAVOOVHANTNSNEEVRIEYISFGIMSLGLLSLAVTSIPSVSNALN 387
 DB 197 SYRYKLLNWAYOOVQONKEDAWIEHDVNRMEIYVSLGIVGLAILALLAVTSIPSVSDSLT 256
 QY 388 WREFSFIQ 395
 DB 257 WREFHYIQ 264

RESULT 5
 Q9RKR6 PRELIMINARY; PRT; 239 AA.
 AC Q9RKR6;
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DE PUTATIVE OXIDOREDUCTASE.
 GN SC75A.08C.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Seeger K.J., Harris D.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.

RC STRAIN=A3(2);
 RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmid and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 DR EMBL; AL133220; CAB61708.1;
 SQ SEQUENCE 239 AA; 24946 MW; 72B3F159A363EDE6 CRC64;

Query Match 6.7%; Score 156.5; DB 2; Length 239;
 Best Local Similarity 27.9%; Pred. No. 0.00018;
 Matches 61; Conservative 41; Mismatches 72; Indels 45; Gaps 10;

QY 25 TKDARKVTGVI-GSGDFAKSLTIRLCGYHVITGSR---NPKFASFFPHVVDVTHHE 80
 DB 23 LPDVSGLVGVGGTGPQCKGLAYRLAKAGQKIVVGSRAAERAAAAAEIGHGVEGADNA 82
 QY 81 DALTKTNIIFVAIHRE-HYTSIWLDRHLVGLKILIDVSNMREINQ-----YPE--SNAE 131
 DB 83 EFARSDVVIIVAVPMDGKGKTLSESLRAELSGKLVVDCVNPGLGDKKAYALKEPESGAAE 142
 QY 132 YLASLPDSLIIVKGFNVVSAAWALQ-----LGPKDASRQVYICSNIIQARQOV 178
 DB 143 QAAALLPDSRVAAAFHLSAVLLQDPEIDETDVMVLGEERADVEI-----VQA---- 192

QY 179 IELARQLNFIPI---IDLGSLSAREIENLPLRLFTLMR 213
 DB 193 --LAGR---IPCMRGVFAGRLRNAHQVESLVANLISVNR 226

RESULT 6
 Q29059 PRELIMINARY; PRT; 213 AA.
 AC Q29059;
 DT 01-JAN-1998 (TRENBLrel. 05, Created)
 DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
 DT 01-AUG-1998 (TRENBLrel. 07, Last annotation update)
 DE CONSERVED HYPOTHETICAL PROTEIN.
 GN AFI209
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
 OC Archaeoglobus
 OX NCBI_TaxID=2234;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Richardson D.L., Kerlavage A.R., Gwinn M., Hickey E.K., Peterson J.D.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
 RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 reducing archaeon Archaeoglobus fulgidus.";
 RL Nature 390:364-370(1997).
 DR EMBL; AE001021; AAB90038.1;
 DR TIGR; AF1209;
 KW Hypothetical protein.
 SQ SEQUENCE 213 AA; 22994 MW; D0D7FD04A6A6AB05 CRC64;

```

Qy 201 IENLPLRLFTIWR 213
      :|| : :|
Db 188 VESITPLILNMR 200

RESULT 8
Q59661 PRELIMINARY; PRT; 224 AA.
ID AC Q59661;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE F420-DEPENDENT NADP REDUCTASE.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC Methanobacterium.
OX NCBI_TaxID=2166;
RN [1]
RS SEQUENCE FROM N.A.
RC STRAIN=MARBURG;
RA Berk H., Thauer R.K.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y17210; CAA76687.1;
SQ SEQUENCE 224 AA; 23448 MW; 8B3514EC14BCBOCF CRC64;

Query Match 6.0%; Score 142; DB 1; Length 224;
Best Local Similarity 29.2%; Pred. No. 0.0022;
Matches 59; Conservative 29; Mismatches 84; Indels 30; Gaps

Qy 37 GSGDFAKSLIRLCGYHVVIGSRNPKFASFFPHVVDVTHHEDALTK--TN----- 87
      :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 8 GTGDGGLGLRLALAGEEVIIGSRDAEKAVSAQKVLIEAERDDLKVKGATNAEAAEA 67

Qy 88 ---IIFVAIHREHYTSLWDLRHLVLGKILDV-----SNNMRINQYPSNAEYLA 134
      :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 68 EVAILTVPLQAQMAT-LGSKVEAIKGVKLIDATVPIDSCILGSGSAVRVIDLWDGSAARAA 126

Qy 135 SLFPD--SLIVKGFNVVSNAALQ--LGPKASRQVYICSNIIARQOVIELAQNLFI-P 189
      :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 127 RFLEQGTTRVAAPFNNSISALLDITGPVDC--DCLIASDRHALDLASELAEKIDGVRA 184

Qy 190 IDLGSLSASAREIENLPLRLFTL 211
      :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 185 IDCGGLENARVIEKITPLINL 206

RESULT 9
Q9UN78 PRELIMINARY; PRT; 217 AA.
ID AC Q9UN78;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE PUTATIVE OXIDOREDUCTASE.
GN WMYQ.
OS Streptomyces coelicolor.
OC Plasmid pSCPI.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycetaceae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1902;
RN [1]
RS SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Bruton C.J., Wietzorrek A., Hartley N., Woodburn L., Chater K.F.;
RL "Genes involved in methylglucosamin biosynthesis from plasmid SCPI of
RT Streptomyces coelicolor A3(2)".
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RS SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RL MEDLINE=88112873; PubMed=2828187;
RX

```

RA Neal R.J., Chater K.F.;
 RT "Nucleotide sequence analysis reveals similarities between proteins
 determining methylenomycin A resistance in Streptomyces and
 tetracycline resistance in eubacteria.";
 RL Gene 58:229-241(1987).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RX MEDLINE=85284984; PubMed=2992952;

RA Chater K.F., Bruton C.J.;
 RT "Resistance, regulatory and production genes for the antibiotic
 methylenomycin are clustered.";
 RL EMBO J. 4:1893-1897(1985).
 DR EMBL; AJ276673; CAB82879.1; -.
 KW Plasmid.
 SQ SEQUENCE 217 AA; 22856 MW; D25CC02FE90A2BF4 CRC64;

Query Match 5.8%; Score 136.5; DB 2; Length 217;
 Best Local Similarity 21.6%; Pred. No. 0.0054;
 Matches 41; Conservative 40; Mismatches 84; Indels 25; Gaps 6;

QY 33 VGVIGSGDFAKSLTIRLCGYHHVIGSRNPKFASEF---FPHVVDVTHHEDALTKNTII 89
 DB 10 IGVLAGNIGRPIGRHNAVAGHEVAFGSRTPQLTFFAGALRPAARATTLAEAVETSVDL 69
 QY 90 FVAIHREHYTSLW-----LRHLVKGKILIDVSNMNRINQYPE-----SNAEYLAS 135
 DB 70 LLSVP-----YSAWDDLQAVGRLAGKLVIDATNPMGLSEGRITSLDQGITQGRNSAK 125
 QY 136 LFPDSLIVKGFNVVS---AWALQGPDKASQVYTCNSNIQARQQVIELARQLNFIPIDL 192
 DB 126 LPEATVVRFAFTHIMDELLWSRGTOQKHPWGMAY-AGDDVDAAKEVVYARLIHDGSPVDL 184
 QY 193 GILSSAREIE 202
 DB 185 GGLDSSAALD 194

RESULT 10
 OS0637
 ID OS0637 PRELIMINARY; PRT; 216 AA.
 AC OS0637
 DT 01-JUN-1998 (TRENBLrel. 06, Created)
 DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
 DT 01-JUN-1998 (TRENBLrel. 06, Last annotation update)
 DE HYPOTHETICAL 23.2 KDA PROTEIN.
 OS Leptosira interrogans.
 OC Bacteria; Spirochaetales; Leptosiraceae; Leptosira.
 OX NCBI_TaxID=173;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ICTERO NO.1;
 RA Takahashi Y., Akase K., Hirano H., Fukunaga M.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DDJB databases.
 DR EMBL; AB010203; BAA24368.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 216 AA; 23176 MW; E7723423AD6EB85C CRC64;

Query Match 5.1%; Score 119.5; DB 2; Length 216;
 Best Local Similarity 23.7%; Pred. No. 0.11;
 Matches 47; Conservative 37; Mismatches 91; Indels 23; Gaps 6;

QY 33 VGVIGSGDFAKSLTIRLCGYHHVIGSRNPKFASEFPHV---VDVTHHEDALTKNTII 89
 DB 6 IGLGSGIVGQTLANGFLKYGAEVKIGTRDFGLKDLAKAGAGASIGSFSAANFGEII 65
 QY 90 FVAIHREHYTSLWDLRHL--LVGKILIDVSNMNR-----INQYPSNAEYLA 136
 DB 66 VLCSKGSVASEVLTSLGIDSLNGKTIIDTNPISIPPONGVNLNFTSYNESLMEKLQK 125
 QY 137 FPDLSLVKGFNVVS AWALQGP--KDASQVYTCNSNIQARQQVIELARQLNFIPIDLGS 194

DB 126 APKANFVKCFSSVGS-GLMVNPQLKGEKPSMFICGNDSSKKQIKRILDTFGWDTEDMGK 184
 QY 195 LSSAREIENLPLRLFTLW 212
 DB 185 VEAAARAIEPLCI---LW 198
 RESULT 11
 OS06CP1
 ID OS06CP1 PRELIMINARY; PRT; 242 AA.
 AC OS06CP1
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
 DE HYPOTHETICAL PROTEIN XF1737.
 GN XF1737.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 OC Xylella.
 OX NCBI_TaxID=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9A5C;
 RX MEDLINE=20365717; PubMed=10910347;

RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorfy H.,
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lompais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.V., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmeri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
 RT "The genome sequence of the plant pathogen Xylella fastidiosa";
 RL Nature 406:151-157(2000).
 DR EMBL; AF003997; AAF84546.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 242 AA; 25385 MW; 857F70152F2B0C1B CRC64;

Query Match 5.0%; Score 116.5; DB 2; Length 242;
 Best Local Similarity 24.0%; Pred. No. 0.21;
 Matches 46; Conservative 40; Mismatches 85; Indels 21; Gaps 7;

QY 28 ARKVTGVIGSGDFAKSLTIRLCGYHHVIGSRNP---KFASEFFPHVDVTHHEDAL 83
 DB 30 AAPMRIGVICASLGTGTVGLVWVKGHEVWFSSRNPKLEAMARELEPR-ASVQQLAAT 88
 QY 84 TKTNIIFVAIHREHYTSLW-DLRHLVKGKILIDVSNMNRINQ---YPESN---AEYLAS 135
 DB 89 EFCVTLLAVPEALPQVGRDLRSAYRGKIVLDSTNPGWASSADVYREARELGVAQTVVK 148
 QY 136 LFPDSLIVKGFNVVS AWALQGPDKASQ-----VYTCNSNIQARQQVIELARQLNFIPI 190
 DB 149 YMPGARLVRAFSAVDATVVE---TSASRRGRIGMPLASDAAEMKVAEGLVRDAGCDPV 205

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlike C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleissmann W.,
 RA Foster C., Gabriellian A.C., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris D., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.:
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003795; AAF57539.1;
 DR FLYBASE; FBgn0034456; CG15121.
 SQ SEQUENCE 408 AA; 47829 MW; 364769862D735B33 CRC64;

Query Match 4.8%; Score 112; DB 5; Length 408;
 Best Local Similarity 19.8%; Pred. No. 0.94; 137; Indels 112; Gaps 17;
 Matches 76; Conservative 59; Mismatches 137;
 QY 132 YIASLFPDSL-----IVKGFNVSAWALQGLGPKDASQVYICSNQINQAOQVI 179
 DB 51 VHYQLFDSLSPKKSVEQDIISCKYNLSLHGVIIRPEET-----SDPFNATQHSY 103
 QY 180 ELAROLNIPIDGLSLSAREIENLPRLFTLWRGP-----VVVAISLATFFELY----- 229
 DB 104 PLEMTNCVMVPLA-----PELPKMYMWV--PLGKYIWTCLFLGTFYVALLRYVH 153
 QY 230 -----SEVRDVIHPYARNOQSDFYKIPTEIVNKTLPITVAITLLSLVLAGLAAAY 280
 DB 154 WREPONATRSYTRNVLHAMALLMFSANNMVKL--KHASIRVIITYTLTYIFGFIITNY 211
 QY 281 QLYYGTKYRRFPPLW---ETWLQCRKQLGLLSFFFAFAMVHVAYSLCLPMRRSERYFLNMA 337
 DB 212 HLSHTAFDMKVPFLRPIDTWSD-----LIHSRLRIVIHDSLLEELRWLPV- 257
 QY 338 YQOVHAN-----INSW---NEEVWRIEMY-----ISFGIMSLGLSLAVTSTP 380
 DB 258 YQALLASPSRSYAVVTQDAFFNRQOKVLIQPIPHLSKVCFG---GLFNALPMASNA 313
 QY 381 SVSNALN-----WREFSF-IOSTLYGYVALLISTFPHVLIYGWKRAPEEYR 425
 DB 314 SPADSLNKFLNWOAGUNWYEEALAFYAEQAGYKVFELDTYPV-----EPLNLE 364
 QY 426 FYTPPNFVLALVLPISVILDLLQ 449
 DB 365 FTTAVIVLSAGIPISSLAFCLEL 388

RESULT 15

Q12754.

ID Q12754 PRELIMINARY; PRT; 1228 AA.

AC Q12754;

DT 01-NOV-1996 (Tremblrel. 01, Created)

DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-NOV-1996 (Tremblrel. 01, Last annotation update)
 DE HYPOTHETICAL 137.5 KDA PROTEIN (LPA5P).
 GN LPA5W OR YP8132.01 OR YPL012W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
 OC Saccharomycetaceae; Saccharomyces.
 OC NCBI_TaxID=4932;
 RP SEQUENCE FROM N.A.
 RA Hall J., Ahmed A., Bussey H., Fortin N., Friesen J.D., Storms R.K.,
 RA Vo D.H., Wang Y., Winnett E.;
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE OF 220-1228 FROM N.A.
 RC STRAIN=AB972;
 RA Badcock K., Churcher C.;
 RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
 [3]
 RP SEQUENCE OF 220-1228 FROM N.A.
 RC STRAIN=AB972;
 RA Walsh S.V., Bartell B., Rajandream M.A.;
 RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
 [4]
 RP SEQUENCE OF 220-1228 FROM N.A.
 RC STRAIN=AB972;
 RA Badcock K., Bowman S., Churcher C.M., Pearson D., Rajandream M.A.,
 RA Walsh S.V., Bartell B.G.;
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U33335; AAB68093.1;
 DR EMBL: Z48483; CA88374.1;
 DR EMBL: Z71255; CAA95029.1;
 SQ Hypothetical protein.
 SQ SEQUENCE 1228 AA; 137507 MW; 92333BCD53CB095F CRC64;

Query Match 4.7%; Score 111.5; DB 3; Length 1228;
 Best Local Similarity 19.9%; Pred. No. 4.4;
 Matches 111; Conservative 71; Mismatches 152; Indels 225; Gaps 28;
 QY 11 KSLSETCLPNGINGIKDKAK-----VTGVIGSGDFAKSLTIRLI 50
 DB 292 KTMATTISSGLAENKYLRLDTIFALRPSNVDTLLTKSWIAVWIKGMYATHQPLKAL 351
 QY 51 R-----CGYHVYIGSRNPKF---ASEFFPHVDVTHHEDAL----- 83
 DB 352 RKIPGVFHMICY---LASETPEVYQAASQCLISILSESVKDDLLYTPSDEKVKNDV 408
 QY 84 -----TKNTIIFVAIHRHYTSLMDLRHLLVKGKILIDVSNMNRINOYPESNAEYLALF 137
 DB 409 EIIISIAKTFIDFLSIRYSHCS-----REIL--KILVAAFNFRY---RSNPHFLKSL- 456
 QY 138 PDSLIVKGFNVSAWALQGLGPKDASQVYICSNQINQAOQVIELARQLNFI----- 188
 DB 457 -----KIVDTWRVN-----EEQFMDLRNIEELVIGASISAMG 488
 QY 189 -----PIDGLSLSAREIENLPRLFTLWRGPW---VAISLATFFELYSFVRVI 236
 DB 489 PEMILAEAPLNLDNPSER-----PCR---AWLLPLIRDYTKNANLATF-----QNEL 533
 QY 237 HPYARNOQSDFYKIPTEIVN-----KTLP-----IVAITLLSL 269
 DB 534 APYIKSFQSKFDKVPESIQLRVFTIVDQINSTLPRFCLEPMDLRSTDFEASELSL 593
 QY 270 VYL-----AGLLAAYQIYGTG-----YREFP-----PWLETWLOCRK 303
 DB 594 LYSEVELRTTICHALKVLAESNVSYAESESSHNVLQLRFPISAEAKNTEYLSL-----K 648
 QY 304 QLGLLSFFFAFAMVHVAYSLCLPMRRS-----ERYL-----FLNMAYQOVHANIESWNE 351,
 DB 649 STNLLAVLNV---YTQTTPNARSYILTDIOLYKITSKEDLEKTFNNVCGLLKNSME 704
 QY 352 EEVWRIEYISFGIMSLGLLSL--AVTSPISVSNALNMRFSFTOSTLYGVALLISTPH 409

Db 705 E SGNVNK KKKPQTATLLDITCMITVLPVSSYSALFSMF-----LTVNSAD 754
Oy 410 VLIYGWKRAFEEYRYFT 428
Db 755 ALIQ--KRA-----YRIIT 766

Search completed: March 14, 2001, 16:13:41
Job time: 82 sec